

Jörg Houpert

Anne-Conway-Straße 1, 28359 Bremen

Verfahren und Vorrichtung zur Reduzierung zufälliger, kontinuierlicher, instationärer Störungen in Audiosignalen

---

Die Erfindung betrifft ein Verfahren und eine Vorrichtung zur Reduktion von Störungen in Audiosignalen, wobei die Störung eine zufällige, instationäre Störgröße  $n(k)$  darstellt, die zu allen Zeitpunkten  $k$  dem Nutzanteil  $s(k)$  des Audiosignals  $x(k)$  überlagert ist. Derartige Störungen werden im Folgenden als zufällig, kontinuierlich und instationär bezeichnet. Dabei liegen die Audiosignale entweder in diskreter Form vor oder werden aus der Abtastung eines analogen zufällig, kontinuierlich, instationär gestörten Audiosignals gewonnen.

Audiosignale sind oftmals durch zufällige, kontinuierliche stationäre und/oder instationäre Störungen - im folgenden kurz auch als Störgeräusche oder Rauschstörungen bezeichnet - beeinträchtigt, die die Qualität des Signals negativ beeinflussen. Die Reduzierung bzw. Entfernung dieser Störgeräusche erfolgt üblicherweise durch die Filterung des gestörten Audiosignals mittels einer Filterfunktion, bei der das gefilterte Ausgangssignal das geräuschreduzierte bzw. ungestörte Audiosignal möglichst gut approximieren soll. Die Berechnung der Filter

funktion erfolgt dabei unter der Annahme der Stationarität des Störsignals.

Im Rahmen der vorliegenden Patentanmeldung wird von der Annahme ausgegangen, dass sich das zufällig, kontinuierlich und instationär gestörte, diskrete Audiosignal  $x(k)$ , welches aus der Abtastung eines analogen, gestörten Audiosignals  $x(t)$  an den diskreten Abtastzeitpunkten  $k$  unter Beachtung des Nyquist-Theorems hervorgegangen ist, additiv aus einem diskreten, ungestörten Audiosignal  $s(k)$ , dem Nutzanteil des Audiosignals, und einem diskreten, zufälligen, kontinuierlichen Störsignal  $n(k)$ , dem Störanteil des Audiosignals, zusammensetzt, wobei  $n(k)$  stationäre und instationäre Rauschanteile umfassen kann:

$$x(k) = s(k) + n(k) \quad (1)$$

Ein bekanntes Verfahren zur Reduzierung bzw. Entfernung derartiger, zufälliger, kontinuierlicher Störungen, das sogenannte Verfahren zur „kurzzeitigen spektralen Dämpfung“ – im folgenden kurz Short Time Spectral Attenuation (STSA) genannt – ist in dem Blockschaltbild gemäß Figur 1 wiedergegeben. Dargestellt ist die Bearbeitung eines Audiosignals  $x(k)$ , welches als Abtastsignal  $x(k)$  des analogen, gestörten Audiosignals  $x(t)$  an den Abtastzeitpunkten  $k$  gewonnen wird.

$X(m, l)$ ,  $S(m, l)$  und  $N(m, l)$  sind die den diskreten Signalen  $x(k)$ ,  $s(k)$  und  $n(k)$  entsprechenden Funktionen beispielsweise im Frequenzbereich, wobei  $m$  die diskrete Frequenz bedeutet. Alternativ kann  $m$  jedoch ein anderer Parameter sein, der eine äquivalente Beschreibung der diskreten Zeitsignale  $x(k)$ ,  $s(k)$  und  $n(k)$  ermöglicht.  $l$  ist die diskrete Zeit des jeweils betrachteten Signalblockes bei der üblichen blockweisen Signalverarbeitung. Im Frequenzbereich gilt daher entsprechend

$$X(m, l) = S(m, l) + N(m, l) \quad (2)$$

Das diskrete Audiosignal  $x(k)$  wird bei diesem bekannten Verfahren in einem ersten Schritt, mittels einer diskreten Fouriertransformation in den Frequenzbereich transformiert, Block 1, so dass sich die diskrete Frequenzbereichsdarstel-

lung  $X(m, l)$  ergibt. Aus dieser diskreten Spektraldarstellung erfolgt bei dem dargestellten Stand der Technik eine einzige und somit stationäre Schätzung  $\hat{\Phi}_{NN}(m)$  der diskreten Autorausleistungsdichte  $\Phi_{NN}(m)$  durch ein bekanntes Schätzverfahren, Block 2, bei dem beispielsweise eine

- (3a) Schätzung der Autorausleistungsdichte innerhalb (annähernd) nutzsingnalfreier Passagen des gestörten Signals oder eine
- (3b) sog. direkte Schätzung

vorgenommen wird.

Die geschätzte diskrete Autorausleistungsdichte  $\hat{\Phi}_{NN}(m)$  ergibt sich aus einem diskreten, zufällig, kontinuierlich gestörten Audiosignal entsprechend des in (3a) genannten Verfahrens durch Auswertung annähernd audiosingnalfreier Passagen des gestörten Signals, in denen näherungsweise gilt

$$x(k) \approx n(k), \text{ da } s(k) \approx 0. \quad (3)$$

Unter Ausnutzung der Linearität der Fouriertransformation ergibt sich innerhalb dieser Abschnitte, in denen  $s(k) \approx 0$ , eine Schätzung der diskreten Autorausleistungsdichte gemäß

$$\hat{\Phi}_{NN}(m) = \Phi_{xx}(m). \quad (4)$$

Hier bedeutet  $\Phi_{xx}(m)$  die Autorausleistungsdichte des gestörten Audiosignals.

Das alternative "direkte Schätzung" genannte Verfahren (3b) wurde in „Steven L. Gay, Jacob Benesty: *Acoustic Signal Processing for Telecommunication* ; Kluwer International Series in Engineering and Computer Science; Kapitel 9: Eric J. Dithorn: *Subband Noise Reduction Methods for Speech Enhancement*, März 2000, ISBN 0-7923-7814-8" vorgestellt und beruht auf einer beschränkten Verfolgung der Leistungsdichte des gestörten Signals.

Basierend auf der Schätzung der Autorausleistungsdichte  $\hat{\Phi}_{NN}(m)$  und der diskreten Frequenzbereichsdarstellung  $X(m,l)$  des diskreten Audiosignals  $x(k)$  erfolgt bei diesem bekannten Verfahren die Bestimmung einer geeigneten Filterfunktion  $H_G(m,l)$ , vgl. Block 3, bei der das abgegebene Signal das ungestörte Audiosignal  $s(k)$  möglichst genau approximiert. In diesem Zusammenhang sind verschiedene Berechnungsvorschriften zur Gewinnung der Filterfunktion  $H_G(m,l)$  bekannt, beispielsweise der

- (6a) Ansatz nach Wiener, bei dem als Approximationskriterium der mittlere quadratische Fehler zwischen Nutzsignal und Schätzung verwendet wird oder der
- (6b) Ansatz zur Amplituden-Subtraktion oder der
- (6c) Ansatz zur Leistungs-Subtraktion,

die in „S. F. Boll; *Suppression of acoustic noise in speech using spectral subtraction*; IEEE Trans. Acoust., Speech & Signal Process.; ASSP-27; Seite 113-120; 1979“, außerdem auch in dem Lehrbuch von P. Vary, U. Heute & W. Hess, „*Digitale Sprachsignalverarbeitung*“, Teubner Verlag, Stuttgart 1998, ISBN 3-519-06165-1, Seiten 380-390, beschrieben sind.

Zur Bestimmung einer Schätzung  $\hat{s}(k)$  des diskreten, ungestörten Nutzteils  $s(k)$  erfolgt eine Filterung des diskreten Audiosignals  $x(k)$  mit der zuvor bestimmten Filterfunktion. Diese kann entweder im Zeitbereich durch die Faltung des diskreten, gestörten Signals  $x(k)$  mit der diskreten Impulsantwort der Filterfunktion  $h_G(k)$  realisiert werden

$$\hat{s}(k) = h_G(k) * x(k), \quad (5)$$

wobei  $*$  den Faltungsoperator darstellt oder wie in Fig. 1 dargestellt, im Frequenzbereich durch die Multiplikation der diskreten Übertragungsfunktion  $H_G(m,l)$  mit der diskreten Spektraldarstellung  $X(m,l)$  des diskreten, gestörten Audiosignals  $x(k,l)$ , vgl. Block 4,

$$\hat{S}(m, l) = H_G(m, l) \cdot X(m, l). \quad (6)$$

Aus der derart bestimmten diskreten Schätzung  $\hat{S}(m, l)$  wird die entsprechende Darstellung  $\hat{s}(k)$  im Zeitbereich durch die inverse, diskrete Fouriertransformation gewonnen, vgl. Block 5, so dass das geräuschbefreite Signal ggf. mittels eines Digital-Analog-Wandlers in ein analoges, geräuschbefreites Signal umgesetzt werden kann.

Ein Nachteil dieses bekannten Verfahrens besteht darin, dass durch die Filterung des gestörten Audiosignals erneut Störungen in das geräuschbefreite Signal eingetragen werden, die durch die Filterung entstehen und zu unerwünschten sog. „Musical Tones“ führen.

Ferner ist aus „M. Berouti, R. Schwartz & J. Makhoul; *Enhancement of speech corrupted by acoustic noise*; in Proc. IEEE ICASSP; Seite 208-211; Washington DC; 1979“ ein weiteres Verfahren bekannt, das im folgenden anhand des Blockschaltbildes gemäß Figur 2 beschrieben ist und welches in seinem Grundprinzip dem Verfahren gemäß Figur 1 entspricht. Dieses bekannte Verfahren arbeitet auf folgende Weise:

Aus einer einmaligen und somit stationären Schätzung der Autorausleistungsdichte  $\hat{\Phi}_{NN}(m)$ , Block 2, und der diskreten Spektraldarstellung  $X(m, l)$ , am Ausgang des Blocks 1 des diskreten Audiosignals  $x(k)$  wird die Filterfunktion  $H_G(m, l)$  ermittelt, Block 3. Vor der eigentlichen Filterung des gestörten Signals, Block 4, wird die Filterfunktion  $H_G(m, l)$  auf einen konstanten, frei gewählten minimalen Wert  $\gamma_{sf}(m)$  - auch als „spektraler Boden“ bezeichnet -, d. h. auf eine maximale Geräuschreduktion, beschränkt, Block 6. Somit ergibt sich für die Filterung eine neue diskrete Filterfunktion  $H_G(m, l, \gamma_{sf}(m))$ , für die gilt:

$$H_G(m, l, \gamma_{sf}(m)) = \begin{cases} H_G(m, l) & \text{für } H_G(m, l) > \gamma_{sf}(m) \\ \gamma_{sf}(m) & \text{sonst} \end{cases} \quad (7)$$

Diese beschränkte Filterfunktion führt einerseits dazu, dass keine Geräuschbefreiung, sondern nur eine Reduktion der Störung möglich ist, andererseits wird die Entstehung von sog. Musical Tones deutlich reduziert.

Das durch die Filterung, Block 4, gewonnene diskrete, geräuschreduzierte Signalspektrum  $\hat{S}(m, l)$  wird anschließend wie in dem Verfahren gemäß Figur 1 durch eine inverse, diskrete Fouriertransformation, Block 5, in den Zeitbereich zurücktransformiert.

Bei beiden bekannten Verfahren erweist es sich als nachteilig, dass diese nur zur Entfernung bzw. Reduktion zufälliger, kontinuierlicher stationärer und ggf. zufälliger, kontinuierlicher langsam instationärer Störungen herangezogen werden können. Zeitliche Änderungen der statistischen Eigenschaften der diskreten Störung  $n(k)$  können nicht oder nur bei sehr langsamen Änderungen erfasst werden. Handelt es sich bei der überlagerten Störung jedoch beispielsweise um ein instationäres Rauschen, ergibt sich eine fehlerbehaftete Schätzung der Autorausleistungsdichte. Hieraus resultiert eine fehlerhafte Bestimmung der Filterfunktion und somit eine Geräuschreduktion, die entweder das eigentliche ungestörte Signal  $s(k)$  beeinträchtigt und/oder das Störsignal  $n(k)$  nur unzureichend reduziert.

Bei Verwendung einer einmaligen und somit stationären Schätzung der Autorausleistungsdichte innerhalb nutzsignalfreier Abschnitte ergibt sich eine fehlerhafte Autorausleistungsdichte, da ein zufällig, kontinuierlich gestörtes Audiosignal in der Regel nicht ausreichend viele nutzsignalfreie Abschnitte aufweist, die eine fortlaufende Aktualisierung der Schätzung der Autorausleistung erlauben. Somit kann der ermittelte Schätzwert den zeitlichen Änderungen der statistischen Eigenschaften des Rauschens keine Rechnung tragen. Zwar wird bei der geschilderten und bekannten „direkten Schätzung“ die Autorausleistungsdichte ständig aktualisiert, jedoch ist die Schätzung fehlerhaft bezüglich des instationären Rauschanteils, wie die diesbezüglichen Betrachtungen in „J. Meyer, K.U. Simmer and K.D. Kammeyer: *Comparison of One- and Two-Channel Noise-Estimation Techniques* ; Proc. 5th International Workshop on Acoustic Echo and

Noise Control (IWAENC-97), vol. 1, Seite 17-20, London, U.K., 11-12 September 1997" zeigen.

Aus dem US-Patent 5,852,567 ist ein weiteres Verfahren zur Reduzierung zufälliger, kontinuierlicher Störungen bekannt. Basierend auf einer Zeit-Frequenz-Transformation wird bei diesem Verfahren versucht, den Signal-Rauschabstand und die Charakteristika des instationären Nutzsignals zu verbessern. Wie bei den zuvor beschriebenen Verfahren erweist es sich auch bei diesem Verfahren als nachteilig, dass es ebenfalls entsprechend seines Entwicklungsziels nur zur Reduzierung von zufälligen, kontinuierlichen, stationären Störungen, nicht jedoch zur Reduzierung von zufälligen, kontinuierlichen, instationären Störungen herangezogen werden kann.

Der Erfindung liegt daher die Aufgabe zugrunde, ein Verfahren und eine Vorrichtung zur Reduzierung zufälliger, kontinuierlicher, instationärer Störungen in Audiosignalen anzugeben, mit dem Ziel, den instationären Rauschanteil im Audiosignal im Verhältnis zu dessen stationären Rauschanteil zu verringern.

Diese Aufgabe wird gelöst durch ein Verfahren gemäß Anspruch 1. Diese Aufgabe wird ferner gelöst durch eine Vorrichtung gemäß Anspruch 11.

Die Vorteile des erfindungsgemäßen Verfahrens und der erfindungsgemäßen Vorrichtung bestehen darin, dass eine Repräsentation des gestörten Audiosignals derart bearbeitet wird, dass die zeitlichen Änderungen der statistischen Eigenschaften des Störanteils des bearbeiteten Audiosignals im Vergleich zum Störanteil des unbearbeiteten Audiosignals verringert werden. Es werden die zeitlichen Änderungen der statistischen Eigenschaften reduziert, so dass das Audiosignal nach der Bearbeitung nur noch durch eine zufällige, kontinuierliche, stationäre Reststörung beeinträchtigt wird und gegebenenfalls eine weitere Reduzierung des durchschnittlichen Rauschpegels zusätzlich vorgenommen werden kann. Bei der Bestimmung der Filterfunktion werden die aktuellen Eigenschaften des Nutz- und des Störsignalanteils berücksichtigt. Das Maß der Geräuschreduktion, also die Filterfunktion, wird nicht auf einen festen Amplitudenwert beschränkt, sondern durch eine Repräsentation der Störung oder einer di-

rekt oder indirekt hieraus ableitbaren Größe dynamisch an die aktuellen zeitlich veränderlichen Eigenschaften des Störsignals angepasst.

Gemäß einer besonders bevorzugten Ausführungsform der Erfindung läßt sich eine Repräsentation der Störung bestimmen, die die zeitlichen Veränderungen der instationären statistischen Eigenschaften des Rauschens beschreibt.

Ein weiterer entscheidender Vorteil des erfindungsgemäßen Verfahrens ist die Einbeziehung der aktuellen Störsignaleigenschaften. Bisherige Verfahren berücksichtigen in diesem Zusammenhang nur einen zeitlich begrenzten Signalausschnitt, so dass sich den ändernden Eigenschaften des Störsignalanteils keine Rechnung getragen wurde.

Vorteilhafte Weiterbildungen der Erfindung sind durch die Merkmale der Unteransprüche gekennzeichnet.

Im Folgenden werden Ausführungsbeispiele der Erfindung anhand der Zeichnung näher erläutert. Es zeigen:

- Figur 1 ein Blockschaltbild eines bekannten Verfahrens zur Reduzierung zufälliger kontinuierlicher Störungen in Audiosignalen;
- Figur 2 ein Blockschaltbild eines weiteren bekannten Verfahrens zur Reduzierung zufälliger kontinuierlicher Störungen in Audiosignalen;
- Figur 3 eine schematische Darstellung des erfindungsgemäßen Verfahrens;
- Figur 4 ein Blockschaltbild einer ersten Ausführungsform des erfindungsgemäßen Verfahrens;
- Figur 5 ein Blockschaltbild einer zweiten Ausführungsform des erfindungsgemäßen Verfahrens;
- Figur 6 ein Blockschaltbild einer dritten Ausführungsform des erfindungsgemäßen Verfahrens,



Figur 7a,

bis 7c der typische zeitliche Verlauf des Störanteiles a) eines gestörten Audiosignales, b) des gemäß dem Stand der Technik bearbeiteten Audiosignales und c) des mit dem erfindungsgemäßen Verfahren bearbeiteten Audiosignales.

Figur 8 eine exemplarische Darstellung der Wirkungsweise des Verfahrens gemäß Fig. 2,

Figur 9 eine schematische Wirkungsweise einer Ausführungsform des bekannten Verfahrens bei Verwendung einer Abschätzung des aktuell enthaltenen Störsignalanteils, welcher die zeitliche Änderung der Störung beschreibt zur Bestimmung der Filterfunktion  $H_G^{dyn}(m, l)$  und deren Beschränkung mittels einer zeitlich konstanten Beschränkungsfunktion  $\gamma_{SF}(m)$

Figur 10 eine exemplarische Darstellung der Wirkungsweise einer Ausführungsform des erfindungsgemäßen Verfahrens.

Figuren 3 und 4 zeigen ein schematisches Blockschaltbild einer ersten Ausführungsform des erfindungsgemäßen Verfahrens. Gemäß dem Blockschaltbild nach Figur 3 wird aus einem diskreten, gestörten Audiosignal  $x(k)$  durch eine entsprechende Transformation, beispielsweise eine Transformation des Signals  $x(k)$  in den Frequenzbereich, eine zugehörige Repräsentation  $X(m, l)$  dieses Audiosignals bestimmt, Block 1. Die Variable  $l$  beschreibt in diesem Zusammenhang den aktuellen Betrachtungszeitpunkt. Diese Repräsentation wird in einer Bearbeitungseinheit 2 bearbeitet. Aus der Bearbeitung dieser Repräsentation entsprechend dem erfindungsgemäßen Verfahren ergibt sich die bearbeitete, neue Repräsentation  $\hat{S}(m, l)$  des Audiosignals, die durch eine Verringerung der zeitlichen Änderungen der statistischen Eigenschaften des enthaltenen Störanteils gekennzeichnet ist. Schließlich wird dann durch eine entsprechende Rücktransformation der diskrete Signalverlauf  $\hat{s}(k)$  gewonnen, der den diskreten, zeit-

lichen Verlauf des geräuschreduzierten Audiosignals als Funktion der diskreten Abtastzeitpunkte beschreibt.

Gemäß Figur 4 wird aus einer Repräsentation des gestörten Audiosignals  $X_s(m, l)$ , - die beispielsweise durch eine geeignete Abbildungsvorschrift aus der Repräsentation  $X(m, l)$  hervorgeht und das aus dem Zeitbereich in den Frequenzbereich transformierte Signal  $x(k)$  repräsentiert-, vgl. Block 5, und der Repräsentation  $\hat{N}(m, l)$ , die eine Schätzung der aktuellen Eigenschaften des Störsignalanteils im Frequenzbereich darstellt, in bekannter Weise unter Heranziehung der Schätzung  $\hat{N}(m, l)$  des Störanteils des Audiosignals eine geeignete Filterfunktion  $H_G^{dyn}(m, l)$  bestimmt. Zusätzlich wird die so bestimmte Filterfunktion  $H_G^{dyn}(m, l)$  unter Heranziehung der Schätzung  $\hat{N}(m, l)$  des Störanteils des Audiosignals dynamisch, d.h. zeitabhängig, beschränkt, vgl. Block 4 und 6. Das Superscript *dyn* kennzeichnet eine Filterfunktion, die durch Einbeziehung der aktuellen Eigenschaften des instationären Störanteils des Audiosignals gewonnen wird.

In einem weiteren Bearbeitungsschritt wird die Repräsentation  $X(m, l)$  des gestörten Audiosignals  $x(k)$  mit der beschränkten Filterfunktion gefiltert, vgl. Block 7, so dass sich ein bearbeitetes diskretes Signal  $\hat{S}(m, l)$  ergibt. Aus dieser Repräsentation  $\hat{S}(m, l)$  ergibt sich mittels einer geeigneten Rücktransformation ein diskreter Signalverlauf  $\hat{s}(k)$ , der dem diskreten, zeitlichen Verlauf des gestörten Audiosignals  $x(k)$  entspricht, jedoch durch eine geringere zeitliche Änderung der statistischen Eigenschaften der enthaltenen Störung gekennzeichnet ist.

Figur 5 zeigt das Blockschaltbild zur Realisierung einer zweiten Ausführungsform des erfindungsgemäßen Verfahrens. Aus dem diskreten, gestörten Audiosignal  $x(k)$  wird jeweils zum Betrachtungszeitpunkt  $l$  beispielsweise durch eine Fouriertransformation eine geeignete Repräsentation  $X(m, l)$  dieses Audiosignals bestimmt, vgl. Block 1. Hieraus wird eine Schätzung  $\hat{N}(m, l)$  des dem ungestör-

ten, diskreten Audiosignal  $s(k)$  überlagerten instationären, zufälligen und kontinuierlichen Störungsanteil  $n(k)$  gewonnen, vgl. Block 4, welche die aktuellen statistischen Eigenschaften des instationären Rauschens beschreibt. Unter Verwendung der Schätzung  $\hat{N}(m, l)$  wird, unter Heranziehung der Repräsentation des gestörten Signals  $X(m, l)$  – die gegebenenfalls durch eine geeignete Abbildungsvorschrift (nicht dargestellt) zusätzlich verändert wird – eine geeignete Filterfunktion  $H_G^{dyn}(m, l)$  bestimmt, vgl. Block 8, die im Gegensatz zu den bekannten Verfahren die Instationarität des Störanteils berücksichtigt. Diese Filterfunktion  $H_G^{dyn}(m, l)$  wird im folgenden Schritt auf einen minimalen Wert  $\gamma_{sf}(m, l)$  beschränkt, vgl. Block 9. Diese Grenze – auch Beschränkungsfunktion genannt – ist nicht konstant, sondern bestimmt sich dynamisch in Abhängigkeit von einer direkten oder indirekten Repräsentation der Störung

$$\gamma_{sf}(m, l) = f(\hat{N}(m, l)). \quad (8)$$

Besonders bevorzugt läßt sich zur Berechnung von  $\gamma_{sf}(m, l)$  zusätzlich auch noch eine Repräsentation des gestörten Audiosignals  $x(k)$  heranziehen. Es gilt dann

$$\gamma_{sf}(m, l) = f(\hat{N}(m, l), X(m, l)). \quad (9)$$

Für die derart beschränkte Filterfunktion  $H_b$  gilt dann:

$$H_b = H_G^{dyn}(m, l, \gamma_{sf}(m, l)) = \begin{cases} H_G^{dyn}(m, l) & \text{für } H_G^{dyn}(m, l) > \gamma_{sf}(m, l) \\ \gamma_{sf}(m, l) & \text{sonst.} \end{cases} \quad (10)$$

Eine geeignete Verknüpfung – beispielsweise eine Multiplikation – einer Repräsentation  $X(m, l)$  des gestörten Audiosignals  $x(k)$  mit der zuvor ermittelten beschränkten Filterfunktion  $H_b = H_G^{dyn}(m, l, \gamma_{sf}(m, l))$  liefert dann ein diskretes Signal  $\hat{S}(m, l)$ , aus dem sich durch eine der Transformation entsprechende Rücktransformation eine diskrete Signalfolge  $\hat{s}(k)$  ableiten lässt, die dem gestörten Audiosignal  $x(k)$  entspricht, jedoch durch eine geringere zeitliche

Audiosignal  $x(k)$  entspricht, jedoch durch eine geringere zeitliche Änderung der statistischen Eigenschaften der enthaltenen Störung gekennzeichnet ist, vgl. Block 6.

In Figur 6 ist ein Blockschaltbild einer dritten Ausführungsform des erfindungsgemäßen Verfahrens dargestellt, welches zur Reduzierung einer zufälligen, kontinuierlichen, instationären Störung in einem Audiosignal dient, das durch eine amplitudenmodulierte Rauschstörung mit konstanter spektraler Färbung beeinträchtigt ist. Aus dem diskreten, gestörten Audiosignal  $x(k)$  wird durch eine schnelle Fouriertransformation (FFT) das diskrete Spektrum  $X(m, l)$  des gestörten Audiosignals im Betrachtungszeitpunkt  $l$  gewonnen, vgl. Block 10.  $X(m, l)$  wird auch als Repräsentationsform des gestörten Audiosignals bezeichnet. Anhand dieses diskreten Spektrums  $X(m, l)$  wird eine Schätzung der zum Betrachtungszeitpunkt  $l$  gültigen Auto-Rauschleistungsdichte,  $\hat{\Phi}_{NN}(m, l)$ , vorgenommen, welche ein Maß für den Störanteil  $n(k)$  im gestörten Audiosignal  $x(k)$  ist. Diese Schätzung erfolgt in zwei Schritten:

- In einem ersten Schritt wird durch eines der bekannten Schätzverfahren ein Schätzwert  $\hat{\Phi}_{NN}(m)$  der stationären Auto-Rauschleistungsdichte ermittelt, die die spektrale Färbung, nicht aber den zeitlichen Verlauf der Störung beschreibt, Block 22.
- In einem zweiten Schritt wird dann eine Größe bestimmt, welche die Instationarität der Störung kennzeichnet, Block 24. Zu diesem Zweck wird aus der geschätzten Auto-Rauschleistungsdichte  $\hat{\Phi}_{NN}(m)$  und dem Spektrum  $X(m, l)$  des gestörten Audiosignals ein zeitvarianter Modulationsfaktor  $\alpha(m, l)$  bestimmt, der die Amplitudenmodulation der Störung beschreibt, beispielsweise

$$\alpha(m, l) = \frac{\min_l |X(m, l)|^2}{\min_m (\hat{\Phi}_{NN}(m))}. \quad (11)$$

Durch Multiplikation der geschätzten, stationären Auto-Rauschleistungsdichte  $\hat{\Phi}_{NN}(m)$  mit diesem Modulationsfaktor ergibt sich dann der gesuchte Schätzwert  $\hat{\Phi}_{NN}(m, l)$  der tatsächlichen Auto-Rauschleistungsdichte  $\Phi_{NN}(m, l)$ , Block 26:

$$\hat{\Phi}_{NN}(m, l) = \alpha(m, l) \cdot \hat{\Phi}_{NN}(m). \quad (12)$$

Hieraus erfolgt unter Einbeziehung der aktuellen, diskreten Fouriertransformierten  $X(m, l)$  des gestörten Audiosignals  $x(k)$  die Bestimmung einer Filterfunktion  $H_G^{dyn}(m, l)$  für den aktuellen Betrachtungszeitpunkt  $l$  mittels eines geeigneten Ansatzes, beispielsweise mittels des bekannten Ansatzes nach Wiener, Block 30.

Die Filterfunktion  $H_G^{dyn}(m, l)$  wird im Folgenden mittels einer dynamisch an die Eigenschaften der Störung angepassten Beschränkungsfunktion,  $\gamma_{SF}(m, l)$  in ihrer Amplitude beschränkt, die sich beispielsweise aus dem zuvor berechneten Modulationsfaktor  $\alpha(m, l)$  entsprechend

$$\gamma_{SF}(m, l) \sim (\alpha(m, l))^\beta \quad (13)$$

mit  $-5 < \beta < +5$ ; besonders bevorzugt ist  $\beta = -1/2$ , proportional verhält, Block 40.

Mittels der auf diese Weise gewonnenen Beschränkungsfunktion lässt sich dann die dynamisch beschränkte Filterfunktion  $H_b$  gemäß Gleichung (10) bestimmen, Block 40.

In einem weiteren Schritt wird anschließend die diskrete Fouriertransformierte des gestörten Signals  $X(m, l)$  mit der zuvor bestimmten, beschränkten Filterfunktion  $H_b$  multipliziert, vgl. Block 50. Aus der sich ergebenden Schätzung  $\hat{S}(m, l)$  lässt sich abschließend durch eine inverse, schnelle Fouriertransformation (IFFT) ein Signal  $\hat{s}(k)$  bestimmen, Block 60, das dem gestörten Audiosignal durch eine verringerte Modulation der Störung, nämlich eine geringere zeitliche Änderung der statistischen Eigenschaften der enthaltenen Störung, und durch

eine von der Beschränkungsfunktion  $\gamma_{sr}(m, l)$  abhängige Geräuschreduktion gekennzeichnet ist.

Figur 7a zeigt den zeitlichen Verlauf eines Störanteils  $n(k)$ , der einem beliebigen diskreten, ungestörten Nutzanteil  $s(k)$  überlagert ist. Wird ein derart zusammengesetztes, diskretes, zufällig, kontinuierlich und instationär gestörtes Audiosignal  $x(k) = s(k) + n(k)$  mittels eines in der Beschreibungseinführung genannten bekannten Verfahrens bearbeitet, so ergibt sich ein Störanteil, der in Figur 7b dargestellt ist. Wird demgegenüber das instationär gestörte Audiosignal  $x(k)$  mit dem erfindungsgemäßen Verfahren bearbeitet, so ergibt sich nach der Bearbeitung der in Figur 7c dargestellte resultierende Störanteil der einen über die Zeit gleichförmigen, stationären Charakter aufweist; die in den Figuren 7a und 7b vorhandene typische Instationarität des Signals ist gemäß Figur 7c erfolgreich beseitigt.

Zur Erklärung der Funktionsweise des erfindungsgemäßen Verfahrens soll im Folgenden von einem Audiosignal  $x(k)$  ausgegangen werden, welches blockweise verarbeitet wird, dessen Repräsentation  $X(m, l)$  dem Quadrat der blockweisen Fouriertransformation entspricht. Das Audiosignal  $x(k)$  soll aus einer instationären Störung  $n(k)$  bzw.  $N(m, l)$  bestehen und kein Nutzsignal  $s(k)$  enthalten. Somit gilt für die diskrete Frequenz  $m_i$  (mit  $i = 1, 2, 3, \dots$ ) und die diskreten Zeitpunkte  $l$ , die den einzelnen Signalblöcken zugeordnet sind:

$$X(m_i, l) = N(m_i, l) \quad (14)$$

Exemplarisch ist in den zugehörigen Abbildungen, Figuren 8a, 9a, 10a der zeitliche Verlauf  $N(m_i, l)$  für eine diskrete Frequenz  $m_i$  wiedergegeben.

Bei Einsatz des bekannten Verfahrens mit begrenzter STSA wird aus der stationären Schätzung der Auto-Rauschleistungsdichte  $\hat{N}(m_i)$  in Teilbild 8a, gestrichelt und dem Störsignal eine Filterfunktion  $H_G$  anhand eines geeigneten Ver-

fahrens (beispielsweise nach Wiener) berechnet, Teilbild 8b. In den Bereichen, in denen die reale Rauschrepräsentation  $N(m_i, l)$  die stationäre Schätzung  $\hat{N}(m_i)$  unterschreitet, nimmt die Filterfunktion  $H_G(m_i, l)$  einen Wert nahe Null an, die Rauschstörung wird zu diesen Zeitpunkten  $l$  annähernd vollständig unterdrückt. Für diejenigen Zeitpunkte  $l$  in denen die Repräsentation der realen Rauschleistungsichte  $N(m_i, l)$  größer als die Schätzung ist, nimmt die Filterfunktion  $H_G(m_i, l)$  hingegen einen Wert nahe Eins an, da ein Teil des aktuellen Rauschsignals als Nutzsignal interpretiert wird.

Begrenzt man diese Filterfunktion gemäß dem STSA-Verfahren auf eine konstante, zeitlich also unveränderliche untere Grenze  $\gamma_{sf}(m_i)$ , so ergibt sich ein zeitlicher Verlauf gemäß Teilbild 8c. Wendet man die so erzeugte Filterfunktion  $H_G(m_i, l, \gamma_{sf}(m_i))$  auf das Störsignal an, so ergibt sich als Ausgangssignal wiederum eine instationäre Reststörung, vgl. Teilbild 8d.

In Figur 9 ist die schematische Wirkungsweise des Verfahrens gemäß Fig. 8 dargestellt, bei dem jedoch die einmalig geschätzte, und somit stationäre Repräsentation der Autorausleistungsichte  $\hat{N}(m_i)$  durch eine dynamische, d.h. durch eine die zeitlichen Änderungen der Störung beschreibende Schätzung der Autorausleistungsichte  $\hat{N}(m_i, l)$  ersetzt wird. Als Filterfunktion  $H_G^{dyn}(m_i, l)$  wird beispielsweise mittels des Ansatzes nach Wiener eine Funktion gewonnen, die durch eine konstante Beschränkungsfunktion  $\gamma_{sf}(m_i)$  gemäß Gleichung (7) auf eine zeitlich unveränderliche Untergrenze festgelegt wird, vgl. Fig. 9c. Erfährt das Filtersignal eine Filterung mit der beschränkten Filterfunktion  $H_G^{dyn}(m_i, l, \gamma_{sf}(m_i))$ , so enthält das bearbeitete Signal gemäß Fig. 9b eine restliche Störung, deren Amplitude gegenüber der Amplitude nach Fig. 8d deutlich reduziert ist, wobei jedoch die Instationarität des Störsignals nicht beseitigt ist.

Ergänzt man das anhand der Fig. 9a bis 9d beschriebene Verfahren um einen weiteren Schritt, so erhält man das erfindungsgemäße Verfahren gemäß Fig. 10. Beschränkt man die Filterfunktion  $H_G^{dyn}(m_i, l)$  gemäß der Figur 9b mittels einer

zeitlich veränderlichen Beschränkungsfunktion  $\gamma_{sf}(m, l)$  beispielsweise gemäß Gleichung (13), so ist es möglich, eine Reststörung im Ausgangssignal zu erreichen, die nahezu oder vollständig stationär ist, die zeitliche Instationarität des Signals  $n(k)$  also nicht mehr enthält. Aus der Schätzung  $\hat{N}(m, l)$ , welche die zeitliche Änderung der Störung beschreibt, Figur 10a, und aus dem gestörten Signal  $X(m, l)$  wird die Filterfunktion  $H_G^{dyn}(m, l)$  bestimmt, vgl. Figur 10b. Diese wird durch eine zeitlich veränderliche Beschränkungsfunktion  $\gamma_{sf}(m, l)$  gemäß Gleichung (10) beschränkt, so dass sich hieraus die dynamisch beschränkte Filterfunktion  $H_b = H_G^{dyn}(m, l, \gamma_{sf}(m, l))$  gemäß den Gleichungen (10) und (13) ergibt, vgl. Figur 10c. Die Filterung des Eingangssignals mit dieser Filterfunktion führt nun zu einem bearbeiteten Signal, das nur noch eine stationäre Reststörung enthält, vgl. Figur 10d.



### Ansprüche

1. Verfahren zur Reduktion zufälliger, kontinuierlicher, instationärer Störungen in Audiosignalen, die in diskreter Form vorliegen oder aus der Abtastung eines analogen zufällig, kontinuierlich, instationär gestörten Audiosignal gewonnen werden, wobei das gestörte Audiosignal mittels einer Filterfunktion gefiltert wird, dadurch gekennzeichnet, dass die Filterfunktion dynamisch unter Berücksichtigung der aktuellen Eigenschaften des gestörten Audiosignals und/oder seiner Bestandteile bestimmt wird, und dass die Filterfunktion dynamisch unter Berücksichtigung der aktuellen Eigenschaften des im gestörten Audiosignals enthaltenen Störanteils begrenzt wird.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet,

dass eine Abschätzung des Störanteils des gestörten Audiosignals erzeugt wird, welche die zeitliche Änderung der Störung beschreibt,

dass aus der Abschätzung des Störanteils die unbeschränkte Filterfunktion  $H_G(m, l)$  in an sich bekannter Weise bestimmt wird,

dass eine Beschränkungsfunktion  $\gamma_{SF}(m, l)$  in Abhängigkeit von dem geschätzten Störanteil des gestörten Audiosignals erzeugt wird, und dass eine beschränkte Filterfunktion  $H_b$  erzeugt wird gemäß

$$H_b = H_G^{dm}(m, l, \gamma_{SF}(m, l)) = \begin{cases} H_G^{dm}(m, l) & \text{für } H_G^{dm}(m, l) > \gamma_{SF}(m, l) \\ \gamma_{SF}(m, l) & \text{sonst.} \end{cases}$$

und zur Filterung des gestörten Audiosignals verwendet wird,

wobei  $m$  die betrachtete diskrete Spektralfrequenz oder ein anderer Parameter ist, der eine äquivalente Darstellung der Signale erlaubt und  $l$  die diskrete Zeit des jeweils betrachteten Signalblockes bei blockweiser Signalverarbeitung ist, wobei ein Block auch nur einen Abtastwert beinhalten kann.

3. Verfahren nach einem der Ansprüche 1 oder 2,

dadurch gekennzeichnet, dass die Beschränkungsfunktion  $\gamma_{sf}(m, l)$  in zeitlicher Abhängigkeit von der zeitlich veränderlichen Schätzung des Störanteils des gestörten Audiosignals erzeugt ist.

4. Verfahren nach Anspruch 3,

dadurch gekennzeichnet, dass die Beschränkungsfunktion  $\gamma_{sf}(m, l)$  in Abhängigkeit von der zeitlich veränderlichen, momentanen Störleistung des geschätzten Störanteils des gestörten Audiosignals erzeugt wird.

5. Verfahren nach einem der Ansprüche 1 bis 4,

dadurch gekennzeichnet, dass die beschränkte Filterfunktion in einem Verfahrensschritt erzeugt wird.

6. Verfahren nach einem der vorstehenden Ansprüche,

dadurch gekennzeichnet, dass die Filterung des gestörten Audiosignals im Zeitbereich, im Frequenzbereich oder in einem anderen mathematisch beschreibbaren Signalraum ausgeführt wird.

7. Verfahren nach einem der vorstehenden Ansprüche,

dadurch gekennzeichnet, dass die unbeschränkte Filterfunktion  $H_g^{*sn}(m, l)$  gemäß einem Ansatz nach Wiener bestimmt wird, bei dem als Approximationskriterium der mittlere quadratische Fehler zwischen Nutzsignal und Schätzung verwendet wird.

8. Verfahren nach einem der vorstehenden Ansprüche,

dadurch gekennzeichnet, dass die unbeschränkte Filterfunktion  $H_g^{*sn}(m, l)$  nach dem Verfahren der Amplituden-Subtraktion bestimmt wird.

9. Verfahren nach einem der Ansprüche 1 bis 8,

dadurch gekennzeichnet, dass das gestörte Audiosignal  $x(k)$  in den Frequenzbereich transformiert wird, anschließend der Störanteil  $N(m, l)$  des transformierten gestörten Audiosignals  $X(m, l)$  abgeschätzt wird, die unbeschränkte Filterfunktion  $H_g^{*sn}(m, l)$  und die Beschränkungsfunktion  $\gamma_{sf}(m, l)$  erzeugt wird und

daraus die beschränkte Filterfunktion  $H_b$  gebildet wird, anschließend das transformierte gestörte Audiosignal  $X(m, l)$  mit der beschränkten Filterfunktion  $H_b$  multipliziert und anschließend in den Zeitbereich zurücktransformiert wird.

10. Verfahren nach Anspruch 1,

dadurch gekennzeichnet, dass die Filterfunktion  $H_G^{dyn}(m, l)$  mittels eines bekannten Ansatzes unter Heranziehung einer Schätzung  $\hat{\Phi}_{NN}(m, l)$  der momentanen Autorausleistungsdichte bestimmt wird.

11. Verfahren nach Anspruch 10,

dadurch gekennzeichnet, dass sich die Schätzung  $\hat{\Phi}_{NN}(m, l)$  der momentanen Autorausleistungsdichte aus einer Gewichtung der Schätzung  $\hat{\Phi}_{NN}(m)$  mit einem zeitabhängigen Gewichtungsfaktor  $\alpha(m, l)$  zu

$$\hat{\Phi}_{NN}(m, l) = \alpha(m, l) \cdot \hat{\Phi}_{NN}(m)$$

bestimmt.

12. Verfahren nach Anspruch 11,

dadurch gekennzeichnet, dass sich der Gewichtungsfaktor  $\alpha(m, l)$  gemäß

$$\alpha(m, l) = \frac{\min(|X(m, l)|^2)}{\min(\hat{\Phi}_{NN}(m))}$$

ermittelt wird, wobei  $X(m, l)$  eine Repräsentation des gestörten Audiosignals ist.

13. Verfahren nach Anspruch 11 oder 12,

dadurch gekennzeichnet, dass die dynamische Beschränkungsfunktion  $\gamma_{SF}(m, l)$  bestimmt wird zu

$$\gamma_{SF}(m, l) = (\alpha(m, l))^\beta, \text{ mit } -5 < \beta < 5.$$

14. Verfahren nach Anspruch 13,  
dadurch gekennzeichnet, dass

$$\beta = -1/2.$$

15. Vorrichtung zur Reduktion zufälliger, kontinuierlicher, instationärer Störungen in Audiosignalen, die in diskreter Form vorliegen oder aus der Abtastung eines analogen, zufällig, kontinuierlich, instationär gestörten Audiosignals gewonnen werden, wobei das gestörte Audiosignal mittels einer Filterfunktion gefiltert wird, gekennzeichnet durch eine Einrichtung (4; 22) zur Abschätzung des Störanteils des gestörten Audiosignals, wobei diese Abschätzung die zeitliche Änderung der statistischen Eigenschaften der Störung berücksichtigt, eine Einrichtung (8; 30) zur Erzeugung einer unbeschränkten Filterfunktion  $H_G^{dm}$  in an sich bekannter Weise unter Berücksichtigung der zuvor ermittelten Abschätzung des Störanteils, welche die zeitlichen Änderungen der statistischen Eigenschaften der Störung berücksichtigt, eine Einrichtung (24, 40) zur Erzeugung einer zeitabhängigen Beschränkungsfunktion  $\gamma_{SF}$  in Abhängigkeit von dem geschätzten Störanteil des gestörten Audiosignals, und eine Einrichtung (7; 40), welche aus der unbeschränkten Filterfunktion  $H_G^{dm}$  und der zeitabhängigen Beschränkungsfunktion  $\gamma_{SF}$  eine beschränkte Filterfunktion  $H_b$  erzeugt, und ein Filter (7; 50), welches das gestörte Audiosignal mit der beschränkten Filterfunktion  $H_b$  filtert.

16. Vorrichtung nach Anspruch 15,  
dadurch gekennzeichnet, dass die Einrichtung (9; 40) die beschränkte Filterfunktion  $H_b$  gemäß

$$H_b = H_G^{dm}(m, l, \gamma_{SF}(m, l)) = \begin{cases} H_G^{dm}(m, l) & \text{für } H_G^{dm}(m, l) > \gamma_{SF}(m, l) \\ \gamma_{SF}(m, l) & \text{sonst.} \end{cases}$$

erzeugt.

### Zusammenfassung

Es wird ein Verfahren und eine Vorrichtung zur Reduktion zufälliger, kontinuierlicher, instationärer Störungen in Audiosignalen vorgesehen, wobei das gestörte Audiosignal mittels einer vorgegebenen Filterfunktion gefiltert wird. Die Filterfunktion wird dynamisch unter Berücksichtigung der aktuellen Eigenschaften des gestörten Audiosignals und/oder seiner Bestandteile bestimmt, und die Filterfunktion wird außerdem dynamisch unter Berücksichtigung der aktuellen Eigenschaften des im gestörten Audiosignal enthaltenen Störanteils begrenzt.